THE EFFECT OF DISCRETE EMOTIONS ON MEMORY: INVESTIGATING THE REPLICABILITY AND BOUNDARY CONDITIONS OF THE MNEMONIC ADVANTAGE FOR DISGUST

John T. West

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology and Neuroscience (Cognitive Psychology).

Chapel Hill
2020

Approved by:

Neil W. Mulligan

Kelly S. Giovanello

Joseph B. Hopfinger

© 2020
John T. West
ALL RIGHTS RESERVED

ABSTRACT

John T. West: Are Learners Sensitive to the Mnemonic Effects of Disgust When Predicting their Memory Performance?
(Under the direction of Neil W. Mulligan)

Research has demonstrated that people are more confident in their ability to remember emotional information. However, such research has almost exclusively defined emotion in terms of valence and arousal. Because discrete emotions also affect memory, with disgusting information being better remembered than frightening information, I sought to determine whether participants are sensitive to the effects of discrete emotion when predicting their memory. Participants were more confident in their memory for emotional (frightening and disgusting) images relative to neutral images. However, because Experiment 1a failed to replicate the mnemonic advantage of disgust, subsequent experiments were concerned with testing the replicability of this effect. The disgust advantage was ultimately replicated in an experiment where participants completed a concurrent secondary task at encoding. These results suggest that the mnemonic advantage for disgust may be more likely to manifest under divided attention, perhaps because the mechanisms which mediate disgust memory are relatively automatic.

ACKNOWLEDGEMENTS

I would first like to express my deep gratitude to my graduate mentor, Dr. Neil Mulligan. My experience learning from Neil during my graduate education has been invaluable to my development as a researcher. During my time studying with Neil, I have been fortunate enough to improve in a number of areas in which he so clearly excels, such as my understanding of psychological theory, experimental design, and my writing.

In addition to thanking Neil, I would also like to express my gratitude to the other members of my Master's committee: Dr. Kelly Giovanello and Dr. Joseph Hopfinger. Since my first semester at Carolina, Kelly has been a warm and supporting presence. I am grateful to have gotten to know her and share interests with her over these past years. I would also like to thank Joe for serving on this committee. Joe's insightful questions have encouraged me to think about the subject matter of this thesis in ways I otherwise would not have, and for that I am grateful.

Lastly, I would like to thank my friends and family for their unending support as I pursue my higher education. Lastly, I would like to thank my partner, Irene Regalario. Irene has been there for me during every step of this process, offering interest and support throughout.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DA	Divided Attention
JOL	Judgment of Learning
LDT	Line Discrimination Task

CHAPTER 1: INTRODUCTION

Metamemory refers to the processes by which learners monitor and predict their memory performance. A common paradigm for studying such monitoring is to have participants make *judgments of learning* (JOLs): judgments about the extent to which they believe they will be able to remember information in the future. In addition to examining monitoring, metamemory researchers also investigate the ways in which learners control their study behavior, as when a participant is allowed to choose how long to study information in preparation for a memory test and may decide to allot study time to some items at the expense of others. The need to better understand metamemory is particularly salient in light of research demonstrating that people use metacognitive judgments to guide study behaviors (e.g., Metcalfe, 2009; Metcalfe & Finn, 2008). Thus, understanding the cues that people use to predict their memory performance would allow us to optimize the way that people study and learn information.

A large body of metamemory research thus far has been concerned with identifying the cues that might influence peoples' JOLs. When discussing examples of such metamemorial cues, it is helpful to consider Koriat's (1997) *cue utilization framework*, a taxonomy of cues which differentiates between intrinsic, extrinsic, and mnemonic cues. *Intrinsic* cues are properties inherent to the study material itself, such as associative strength or word frequency. According to the cue utilization framework, people are typically sensitive to such cues and tend to incorporate them into their JOLs. An example of an intrinsic cue is semantic relatedness, or the extent to which words in a pair share semantic overlap. In an associative learning experiment, subjects might be asked to study word pairs which differ in their relatedness, with some word pairs being highly related (e.g., dog-cat) and others being only weakly related (e.g., flower-pencil). Such research has found that greater semantic relatedness is associated with both higher JOLs and higher memory performance, suggesting that learners possess some amount of metacognitive awareness of the memory-enhancing effects of relatedness (Koriat, 1997; Rabinowitz et al., 1982).

Although participants' JOLs are likely to be influenced by such intrinsic cues, not all cues affect metamemory. The cue utilization framework posits the existence of a second class of cues: *extrinsic* cues, or properties of the study context external to the to-be-learned stimuli such as the retention interval or the presence of repeated practice (Koriat, 1997). According to the cue utilization framework, because such cues are not directly observable when participants predict their memory performance, participants tend to discount these cues relative to intrinsic cues. Consistent with this proposition, Shaw and Craik (1989) found that although deeper processing at encoding is associated with marked improvements in memory performance, participants are largely insensitive to this effect when making JOLs. Similarly, although repeated retrieval practice improves memory performance, research has found that practice does not increase JOLs to the same extent (Koriat et al., 2002).

A final category of cues posited by the cue utilization framework is the class of *mnemonic* cues, which reflect the influence of internal, phenomenological experiences during study which might affect learners' JOLs (Koriat, 1997). To the extent that such internal experiences are correlated with memorability, the use of such cues when predicting future memory might result in well-calibrated JOLs. However, research has shown that at times people may mistake phenomenological experiences as indicators of memorability even when these experiences are not associated with better memory. For instance, although people often give higher JOLs to materials which are processed with a greater degree of perceptual fluency, perceptual fluency often does not improve later memory (Rhodes & Castel, 2008, 2009). Strikingly, perceptual fluency can even have opposing effects on predicted and actual memory. For example, Besken and Mulligan (2014) had participants study and give JOLs to spoken words which were either intact (fluent) or which contained inter-spliced silence (disfluent). As predicted, participants gave higher JOLs to perceptually fluent words. The prediction that the intact words would produce the best memory was, however, a metacognitive illusion; it was the *disfluent* words which were better remembered. The existence of such a double dissociation between predicted and actual memory suggests that we sometimes lack proper insight into the factors that affect our memory, a fact which further reinforces the need to better understand metamemory.

Taken together, the aforementioned metamemory research suggests that people use a variety of cues to predict their memory performance. Although these cues often lead people to make predictions which track well with memory performance (e.g., associative strength), variables exist which influence memory performance but not memory predictions (e.g., levels of processing), and which influence predictions but not performance, or which affect predicted and actual memory in opposing ways (e.g., perceptual fluency). Because JOLs are subject to such metacognitive illusions, determining how various cues affect the calibration between JOLs and memory performance is an important challenge for metamemory researchers to overcome.

The present study is concerned with the intrinsic cue of stimulus emotionality. Given that people sometimes find themselves in emotionally-charged situations in which they must monitor the accuracy of their memory (e.g., when one experiences a crime and must determine how confident they are in their memory for the details of this event), basic research on emotion and metamemory carries the benefit of potentially informing applied research in areas such as eyewitness testimony. Before I discuss research which has examined the effect of emotionality on JOLs, I will first briefly review basic research on emotional memory.

In general, laboratory research on emotion and memory has found that emotional information is better remembered than non-emotional information (for reviews, see Bennion et al., 2013; Buchanan, 2007). For example, Kensinger and Corkin (2003) tested participants' recognition memory of emotional and neutral words (Experiment 1). In addition to remembering the emotional words more accurately than the neutral words, participants remembered the emotional words with a greater degree of recollection. Other studies have shown that emotion enhances the memorability of images as well as words. Specifically, Bradley et al. (1992) tested participants' memory for emotional and neutral images using immediate and delayed free recall tests. As expected, emotional images were better remembered than neutral images, both immediately and after a one-year delay.

Researchers interested in explaining this emotional enhancement effect have generated a number of explanations as to why stimulus emotionality might enhance memory performance. Of note, research suggest that the mechanisms which underlie this effect depend on whether memory is tested immediately or at a delay. At longer retention intervals the memory-enhancing effects of emotion have been attributed to the beneficial effects of activity in the basolateral amygdala on memory consolidation (McGaugh, 2004). Indeed, research has shown that the performance advantage for emotional compared to neutral information becomes larger at longer delays, at which point amygdala-mediated consolidation processes have had a chance to act upon participants' memory traces. For example, in a series of experiments Sharot and Phelps (2004) tested participants' recognition memory for emotion and neutral words both immediately and after a 24-hour delay. Their results showed that the effect of emotion and memory performance was moderated by retention interval, such that whereas participants' memory for neutral information become significantly worse over time, participant's memory for emotional information either became better or remained the same.

In contrast, the immediate effects of emotion on memory have typically been explained by a set of cognitive factors. Specifically, Talmi and colleagues have found that the immediate effects of emotion might be explained in terms of the combined effects of relatedness and distinctiveness (Talmi et al., 2007; Talmi & Moscovitch, 2004). Emotional items are often perceived as having greater thematic relatedness due to their shared emotional nature. For example, because the words "death" and "pain" are both aversive, these words might be perceived as more related than the neutral words "chair" and "coffee". In addition to being more interrelated, emotional information is thought to be more distinctive. According to Talmi et al. (2007), because emotional information is more likely to be relevant to currently active goals than neutral information (e.g., an image of a gun is relevant to the goal of survival, whereas an image of a shoe is not; Lazarus, 1991), emotional information possesses the unique feature of goal relevance. When emotional items are presented alongside neutral items in mixed lists, their goal relevance may cause them to stand out, making them relatively more distinctive. Given that both distinctive and related information are better remembered (Hunt & McDaniel, 1993), the increased distinctiveness and relatedness of emotional information might be part of what makes it more memorable.

In order to test whether increased relatedness mediated the effect of emotion on immediate memory, Talmi et al. (2007) conducted a series of experiments in which participants studied three types of images: unrelated emotional images, unrelated neutral images, and highly inter-related neutral images. In order to simultaneously examine the mediating effects of distinctiveness, the authors manipulated whether participants studied pure lists comprised of only a single stimulus type, or mixed lists which were composed of both emotional and neutral images. The rationale for this pure/mixed list manipulation was that although an emotional item might stand out relative to neutral items in a mixed list, this same item would be less distinctive compared to an entire list of similarly emotional items. Talmi et al. found that although relatedness and distinctiveness did not explain the effects of emotion on memory independently, these factors did explain the emotional advantage together; emotional images were not better remembered than inter-related neutral images in pure lists. In sum, the extant research suggests that emotional information is better remembered than neutral information, and that for immediate tests the emotional memory advantage is explained by the higher relatedness and distinctiveness of emotional information, and at a delay these cognitive factors are supplemented by amygdala-mediated consolidation advantages (see Talmi, 2013 for a review).

When considering the effect of emotion on episodic memory, researchers have primarily operationalized emotion in terms of two affective dimensions: valence and arousal. *Valence* refers to the extent to which a stimulus evokes positive or negative affect. *Arousal* refers to the level of excitement or intensity that an emotional stimulus evokes, regardless of its valence. For example, although an image of a funeral and a gruesome operation are both negatively valenced, the latter is more intense (higher in arousal) than the former. Valence and arousal are often confounded in memory research, such that emotional items might not only be significantly more valenced (positively or negatively) than the neutral items they are compared to, but also more arousing. Although it is therefore often difficult to make inferences about the unique effects of valence and arousal on memory performance, some research has specifically addressed this issue. For example, Kensinger and Corkin (2004) disentangled the mechanisms underlying the mnemonic effects of arousal and valence by having participants study neutral words, negative and arousing words (i.e. “rape”), and negative and non-arousing words (i.e. “sorrow”) under full or divided attention. Whereas the memory advantage for negative, non-arousing words compared to neutral words was present in the full attention condition, this benefit disappeared under divided attention. This suggests that the unique effect of valence on memory may be due to controlled, resource-demanding processes such as elaboration. In contrast, the memory advantage for negative, arousing words relative to neutral words was present in both the full and divided attention conditions. This suggests that the effects of arousal are due to more automatic, resource-sparing processes such as automatic attention capture. Such results support the distinction between valence and arousal in memory research.

Despite the large body of research which has examined the effect of stimulus emotionality on memory performance, much less is known about how emotion affects metamemory. What little research exists suggests that – consistent with emotionality’s classification as an intrinsic cue (Koriat, 1997) – emotion does indeed impact metamemorial judgments, with people giving higher JOLs for emotional items compared to neutral items (e.g., Hourihan & Bursey, 2017; Hourihan et al., 2017; Nomi et al., 2013; Tauber & Dunlosky, 2012; Tauber et al., 2017; Witherby & Tauber, 2018; Zimmerman & Kelley, 2010). In the following paragraphs I will review the extant metamemory research which has examined the effect of emotion on JOLs, which has focused mostly on the effects of valence and arousal.

What little emotional metamemory research exists has consistently found that regardless of whether emotion enhances memory performance, people give higher JOLs to items with higher levels of valence (both positive and negative) and arousal. For example, Zimmerman and Kelley (2010) found that participants gave higher JOLs to emotional (both positive and negative) relative to neutral words in preparation for free and cued recall tests.[1] Other researchers have replicated these findings by demonstrating that participants give higher JOLs to emotional words (Hourihan et al., 2017; Tauber & Dunlosky, 2012). In addition to words, researchers have demonstrated that emotional images are given higher JOLs than neutral images (Hourihan & Bursey, 2017; Tauber et al., 2017). Hourihan and Bursey (2017) – for example – presented participants with positive and neutral images and found that, although positive valence did not consistently enhance recognition performance, participants gave higher JOLs to positive compared to neutral images. Lastly, research has demonstrated that participants give higher JOLs to faces displaying emotionally charged facial expressions as well (i.e. smiling, frowning; Nomi et al., 2013; Witherby & Tauber, 2018). As an example, Nomi et al. (2013) had participants study and give JOLs to angry, happy, and neutral facial expressions in preparation for a recognition test and found that participants gave higher JOLs to angry and happy faces compared to neutral faces.

Notably, although many of the aforementioned studies of emotional metamemory have confounded valence and arousal, Hourihan et al. (2017) manipulated valence and arousal independently in order to disentangle the effects of these dimensions. Specifically, the authors varied arousal but not valence in one experiment (Experiment 1), and valence but not arousal in another (Experiment 2). Their results showed that both valence and arousal are independently associated with higher JOLs (but see Tauber et al., 2017). Taken together, this research –

[1] Because the positive and negative words were also more arousing than the neutral words, it is not possible to disentangle the unique effects of valence and arousal on JOLs in this study

although preliminary – suggests that participants are sensitive to valence and arousal when predicting their memory performance.

In order to determine the mechanism by which emotion influences JOLs, Hourihan et al. (2017) contrasted two potential causes of the JOL-enhancing effects of emotion. According to the authors, the influence of emotion on JOLs might be due either to an unconscious process by which participants mistake the physiological sensation of arousal as an indicator of memorability (the "physiological account"), or a conscious process by which participants apply beliefs about the effects of emotion on memory to their performance predictions (the "cognitive account").[2] In order to test these mechanisms, Hourihan et al. had participants study and give JOLs for words which varied in arousal but not valence (Experiment 1) or which varied in valence but not arousal (Experiment 2). In contrast to the physiological account, which predicts a unique effect of arousal but not valence on JOLs, participants gave higher JOLs to both valenced and arousing words. To further compare the physiological and cognitive accounts of emotionally-enhanced JOLs, the authors varied valence and arousal continuously within a single list in order to make each word's emotional status less distinct/noticeable relative to other items in the list (Experiment 3). The authors reasoned that if emotional words are given higher JOLs due to the application of a belief about memory, emotional words should only be given higher JOLs when the emotional nature of such words is obvious to participants as they make JOLs. Consistent with the cognitive account, valence and arousal did not influence JOLs when the items' emotional status was made less obvious. Thus, research suggests that participants give higher JOLs to valenced and arousing stimuli, likely because participants apply a (potentially implicit) belief that emotional items will be better remembered than neutral items.

[2] These explanations are similar to Koriat's (1997) distinction between experience-based and theory-based contributions to JOLs respectively.

An important limitation of emotional metamemory thus far is that this research has almost exclusively defined emotion in terms of valence and arousal. Such research makes the assumption that emotional experience can be sufficiently explained in terms of such continuous properties, a proposition shared by *dimensional models* of emotion (Barrett, 2006; Russell, 1980). According to Russell's (1980) *circumplex model of affect* – for example – emotional experiences can be sufficiently described in terms of the intersection between the continuous dimensions of arousal and valence. Excitement, for example, is an emotional state which combines positive valence with high arousal, whereas depression combines negative valence with low arousal.

Whereas such models characterize emotions according to continuous dimensions, other researchers have taken a different approach. Specifically, *discrete models* of emotion claim that conceptualizing emotion in terms of dimensions such as valence and arousal is an oversimplification. Instead, discrete models posit that qualitative differences between the various discrete categories of emotion (i.e. fear, disgust, sadness, anger) reflect important properties of emotion which are not appreciated by models which reduce emotions to their dimensional components (Ekman, 1992). Indeed, emotional theorists have suggested that various emotions are associated with unique, situationally relevant patterns of motivation and behavior (Frijda, 1987). Fear – for example – is evoked when one encounters a threat to well-being, and motivates one to avoid that threat. In contrast, anger is evoked when one encounters an obstacle that impedes progress towards a currently active goal, and motivates one to confront and remove this obstacle. To the extent that differences between emotional categories are an inherent aspect of emotional experience, it might be predicted that discrete emotional categories should have unique effects on memory above and beyond the effects of valence and arousal (for a similar argument, see Levine & Pizarro, 2004; Levine & Edelstein, 2009).

Evidence in favor of discrete models of emotion as applied to episodic memory can be seen in research on memory for disgusting information. A number of studies using a variety of materials have found that disgusting information is better remembered than frightening information (Chapman, 2018; Chapman et al., 2013; Charash & McKay, 2002; Croucher et al., 2011; Ferré et al., 2018; Zhang et al., 2018), suggesting that disgust has a relatively privileged place in episodic memory compared to the similarly negative, avoidance-oriented emotion fear (but see Marchewka et al., 2016, who did not find a significant difference in recognition between disgusting and frightening images). Notably, researchers have found that even after equating disgusting and frightening images in terms of arousal, valence, subjective distinctiveness, visual complexity, and semantic relatedness, disgusting images are still remembered better than frightening images (Chapman, 2018; Chapman et al., 2013). The fact that the advantage for disgusting information in immediate memory persists even after relatedness and distinctiveness have been equated suggests that the mechanisms which explain the mnemonic advantage for disgusting information might differ from those that explain the advantage for emotion defined more broadly (Talmi et al., 2007; Talmi & Moscovitch, 2004). In addition to being better remembered, evidence suggests that when matched for valence, disgusting images are recognized with a greater degree of recollection than frightening images (Croucher et al., 2011).

In addition to demonstrating this disgust advantage with images, researchers have found that disgusting words are better remembered than frightening words when the two categories are equated for arousal and valence (Ferré et al., 2018). Similarly, Zhang et al. (2018) demonstrated better recognition for disgusted relative to fearful faces. The finding that disgusting information is better remembered than frightening information is difficult to explain from the perspective of dimensional models of emotion, which tend to focus on valence and arousal as the sole causes of emotionally enhanced memory.

In light of research indicating that learners are sensitive to the effects of emotion when making JOLs, as well as research suggesting that the effects of emotion on episodic memory cannot be completely explained by dimensional models of emotion, the present study sought to determine whether learners are sensitive to discrete emotions when predicting their memory performance. More specifically, the present study sought to determine whether the differential effects of disgust and fear on episodic memory performance would generalize to metamemory judgments.

Because discrete models of emotion highlight the importance of distinguishing between the effects of categorically different emotions, these models suggest that discrete emotions such as disgust might differentially affect cognition in general, and memory and metamemory in particular. Despite this possibility, there is preliminary evidence suggesting that people may be insensitive to discrete emotional categories when predicting their memory performance. Specifically, Witherby and Tauber (2018) presented participants with facial expressions displaying angry, sad, fearful, and neutral expressions and asked them to make JOLs in preparation for a recognition test. Whereas participants gave higher JOLs to negative compared to neutral expressions, JOLs did not differ between angry, sad, and fearful expressions.

Although it is therefore possible that learners are not sensitive to discrete emotion when making JOLs, an important limitation of this study renders such a conclusion premature: memory performance did not differ between angry, sad, and fearful expressions. The finding that memory performance did not differ between emotions has two possible explanations. First, it is possible that although stimuli from certain emotional categories such as disgust and fear differ in their memorability, stimuli from other categories like anger and sadness do not. It is also possible that the stimuli in Witherby and Tauber’s (2018) study did not differ in their memorability because the authors used facial expressions rather than words or images. Because images and words are intended to evoke an emotional reaction whereas facial expressions are simply the reflection of an emotional reaction in another person, it is possible that the principles which govern the emotional memory of words and images might differ from those which govern the recognition of emotional expressions.

Due to the lack of a difference in memory between the emotional faces in Witherby and Tauber’s study, it is remains possible that learners are sensitive to discrete emotions when making JOLs, but only when such discrete emotions actually have differential effects on memory. After all, in a situation where discrete emotions are equally memorable, participants are right to assign these categories equivalent JOLs. A stronger test of the effects of discrete emotions on metamemory would be to see whether participants assign different JOLs to emotional categories which actually differ in memorability. Given that disgust has consistently been shown to enhance memory performance relative to fear, it is the ideal emotion to test whether discrete emotions similarly affect metamemory.

CHAPTER 2: EXPERIMENT 1A

The aim of Experiment 1a was to test whether disgusting images are given higher JOLs than equally arousing and negatively valenced frightening images. In doing so, I sought to determine whether discrete or dimensional models of emotion more accurately apply to metamemory. From the perspective of dimensional models of emotion, it might be predicted that only valence and arousal should impact metamemorial judgments, and that after equating frightening and disgusting images for these dimensions, participants will not differ in their JOLs for frightening and disgusting images. Indeed, this outcome would be consistent with theorizing in emotional metamemory research thus far, which has implicated valence and arousal as being central to the effects of emotion on metamemory (see Hourihan et al., 2017; Tauber et al., 2017). If dimensions such as valence and arousal affect metamemory but discrete emotions do not, this might result in a situation in which actual and predicted memory conflict, resulting in a metamemorial illusion. Specifically, if disgusting images are better remembered than frightening images but participants do not give disgusting images higher JOLs, participants may be underconfident in their memory for disgust by failing to take the mnemonic advantage for disgust into account.

However, even after equating disgusting and frightening stimuli for valence and arousal, it is possible that participants may assign higher JOLs to disgusting compared to frightening images. Hourihan et al.'s (2017) results suggest that participants give higher JOLs to emotional items because they notice their emotionality relative to neutral items and apply a belief about memorability to their performance predictions. Although participants in the present experiment will no longer have the cues of valence and arousal, they will have the intrinsic cue of discrete emotion. Thus, to the extent that participants notice and differentiate between the different types of emotions elicited by a disgusting cockroach and a threatening gun – for example – one might predict that participants will incorporate these reactions into their JOLs. Such a result would suggest that the differential effects of emotional categories suggested by discrete models of emotion apply not just to memory performance, but to metamemorial judgments as well.

Method

The design of Experiment 1 was based on a prior study which found the disgust advantage in free recall using frightening and disgusting images equated for valence, arousal, and relatedness (Chapman, 2018; Experiment 1). The design of the present experiment and Chapman's experiment differed only in that 1) participants in our study provided JOLs and therefore engaged in intentional rather than incidental encoding and 2) participants did not complete the line discrimination task which Chapman used to assess attentional biases towards different emotions during encoding. Given that an earlier study Chapman et al. (2013; Experiment 3) found that the disgust advantage in recognition did not depend on whether participants engaged in the line discrimination task during encoding, I did not expect this detail to affect our results. Because no study to our knowledge has examined whether the disgust advantage for image recall is present during intentional encoding, or when participants make JOLs during encoding, it was unclear how these changes might affect the presence of the disgust advantage.

Participants

The procedures in this paper were approved by the Office of Human Research Ethics of the University of North Carolina at Chapel Hill. Our sample size was chosen by using an a priori power analysis based on the size of the main effect of Emotion (3: disgust, fear, neutral) on recall performance reported by Chapman ($\eta^2_p = 0.33$; 2018; Experiment 1). This power analysis indicated that a sample size of 25 participants would be sufficient to achieve 99% power. As such, participants were 25 undergraduate students who were compensated for their participation with course credit.

Design and Materials

The design of Experiment 1a was a 3 (emotion: disgust, fear, neutral) x 2 (memory measure: predicted [JOL] vs. actual memory performance) within-subjects design. The images used were obtained from Chapman (2018; Experiment 1), who obtained images from internet searches, as well as from a prior study demonstrating the disgust advantage (Chapman et al., 2013). The materials consisted of photographs of 14 disgusting (e.g., insects, medical operations), 14 frightening (e.g., dangerous animals, war), and 14 neutral (e.g., coat, basket) stimuli. Pilot data from Chapman's (2018) study indicated that the frightening and disgusting images were equally arousing and negatively valenced. Additionally, Chapman found that although disgusting images were rated as more disgusting than the frightening images and frightening images were rated as more frightening than the disgusting images, the disgusting images were as disgusting as the frightening images were frightening. These data suggest that participants perceived the stimuli as belonging to the appropriate emotional category, and that the two categories did not differ in the degree of discrete emotion that they evoked. Lastly, Chapman calculated relatedness by presenting subjects with pairs of disgusting images and pairs of frightening images and asking them to rate their relatedness using a Likert scale. After quantifying the relatedness of each image by calculating the average relatedness rating for a given image with every other image in its emotional category, Chapman found that the disgusting and frightening images did not differ in their relatedness.[3]

Procedure

Experiment 1 consisted of four phases: a study phase, a 45 min delay, a free recall test, and a stimulus-rating phase. During the study phase, participants were first told that they would study a series of images in preparation for an upcoming memory test which would occur approximately 45 min later. Each trial began with a 500 ms blank screen, after which a random image was presented for 2 s, with image order being randomized anew for each subject. After another 500 ms blank screen, participants were asked to rate how confident they were that they would be able to remember the image in 45 min. JOLs were self-paced and were made on a 0 to 100 scale, with 0 meaning “not confident at all” and 100 meaning “extremely confident”. Participants were encouraged to use the entire scale. After studying and making JOLs for all 42 images, participants completed a series of unrelated tasks for 45 min. Specifically, participants read a series of emotionally-neutral, educational passages related to geology and answered comprehension questions about their content.

After a 45 min delay, participants’ memory was tested using a free recall test. During the recall test, participants were provided with a sheet of lined paper and were instructed to write down a brief description (i.e. a few words) of any image that they could remember regardless of

[3] Because Chapman (2018) only assessed relatedness for disgusting and frightening images during her pilot experiment, it is unclear whether the emotional images used were normatively more related than the neutral images. However, because our primary theoretical interests in the present study have to do with the comparison between the frightening and disgusting images, this issue is not directly relevant to the current investigation.

presentation order. The recall testes lasted a maximum of 10 mins, but participants could terminate the test early if they felt that they could not recall any more images.

Lastly, participants were shown all 42 images once again in a random order and were asked to rate each image in terms of its valence, arousal, disgust, and fear using a series of 1-7 Likert scales. Asking participants to provide such ratings was useful in that doing so allowed me to determine whether the images had indeed been correctly matched for potential confounds such as valence and arousal within our specific sample. Participants were tested individually, and the entire procedure took approximately 75 mins.

Data Analysis

Recall performance was coded by two independent raters (myself and a trained undergraduate research assistant). Every written description was coded as either referring to a specific studied image or as not referring to a studied image (i.e. an intrusion).[4] Consequently, certain descriptions which were too vague to determine which image was being remembered were marked as intrusions (e.g., “something gross”). All disagreements were resolved through discussion.

Of the 404 descriptions written down across 25 participants, there were 18 disagreements between raters (i.e. raters agreed on 95.6% of descriptions). Additionally, subjects’ average recall performance correlated very highly between raters (r = 0.98). Because recall scoring was determined to be highly reliable, in subsequent experiments recall was coded by a single rater only (myself).

Analyses for the following experiments consisted primarily of a series of mixed and repeated-measures ANOVAs. Following the recommendations of Girden (1992), whenever

[4] Intrusions were infrequent across all experiments: [Experiment 1a: (M = 0.44, SD = 0.65), Experiment 1b: (M = 0.64, SD = 0.99), Experiment 2: (M = 0.50, SD = 1.06)].

Mauchly's Test of Sphericity indicated a significant violation of sphericity, the Huynh-Feldt correction was applied for values of epsilon above 0.75 and the Greenhouse-Geisser correction was applied for values of epsilon below 0.75. Post-hoc analyses consisted of pairwise comparisons of means with Bonferroni corrections for multiple comparisons.[5] Within-subjects confidence intervals displayed in figures were constructed according to the methods described by Morey (2008). All analyses were conducted using R version 3.6.2.

Results

Recall and JOLs

I began by conducting a 3 (Emotion: disgusting, frightening, neutral) x 2 (Measure: predicted [JOL] vs. actual memory performance) repeated-measures ANOVA (see Figure 1). Conducting a repeated-measures ANOVA which includes the Measure factor allows one to test for effects of emotion on *calibration*: the extent to which predicted memory matches actual memory. In other words, if emotional categories were differentially associated with overconfidence or underconfidence, this analysis would reveal such effects.

This analysis revealed a main effect of Emotion [$F(2, 48) = 51.53$, $p < 0.001$, $\eta^2_p = 0.682$], reflecting higher predicted and actual memory for emotional relative to neutral images. The nature of this effect is expanded upon in subsequent analyses below. Additionally, there was a main effect of Measure [$F(1, 24) = 13.37$, $p = 0.001$, $\eta^2_p = 0.358$], reflecting the fact that participants were generally overconfident in their predictions. There was – however – no significant interaction between Emotion and Measure ($F = 2.26$, $p = 0.115$). As such, calibration did not vary by emotion.

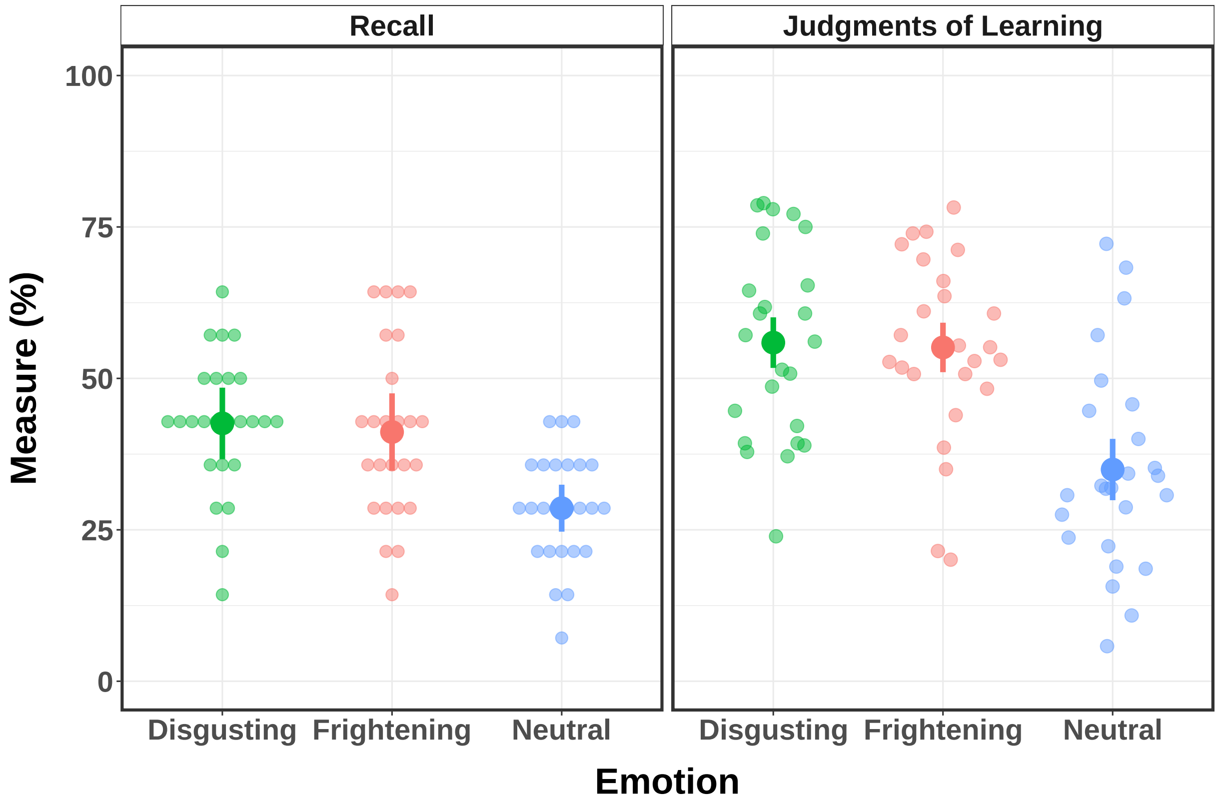
[5] Rather than dividing alpha by the number of comparisons, a mathematically equivalent approach was taken in which p-values were adjusted by multiplying the uncorrected p-values by the number of comparisons and setting their maximum value to 1. As such, all instances in which post-hoc p-values are equal to 1 reflect instances in which the corrected p-values would have exceeded 1 after multiplication.

In order to examine the effects of emotion on recall performance specifically, a repeated-measures ANOVA was conducted with Emotion (3) as a within-subjects factor. There was a significant effect of Emotion on recall performance [$F(2, 48) = 11.15$, $p < 0.001$, $\eta^2_p = 0.317$]. Post-hoc tests revealed that participants recalled fewer neutral images compared to both frightening ($p = 0.002$) and disgusting images ($p < 0.001$). However, participants did not differ in their recall of frightening and disgusting images ($p = 1.000$).

Next, a repeated-measures ANOVA with Emotion (3) as a within-subjects factor was conducted on JOLs. Similar to the analysis of recall performance, there was a significant effect of Emotion on JOLs [$F(2, 48) = 61.07$, $p < 0.001$, $\eta^2_p = 0.718$]. Post-hoc tests revealed that neutral images were given lower JOLs than both frightening and disgusting images (both p's < 0.001). However, participants did not differ in their JOLs for disgusting and frightening images ($p = 1.000$).

Figure 1

Recall and Judgments of Learning by Emotion in Experiment 1a



Note. Error bars represent 95% within-subject confidence intervals.

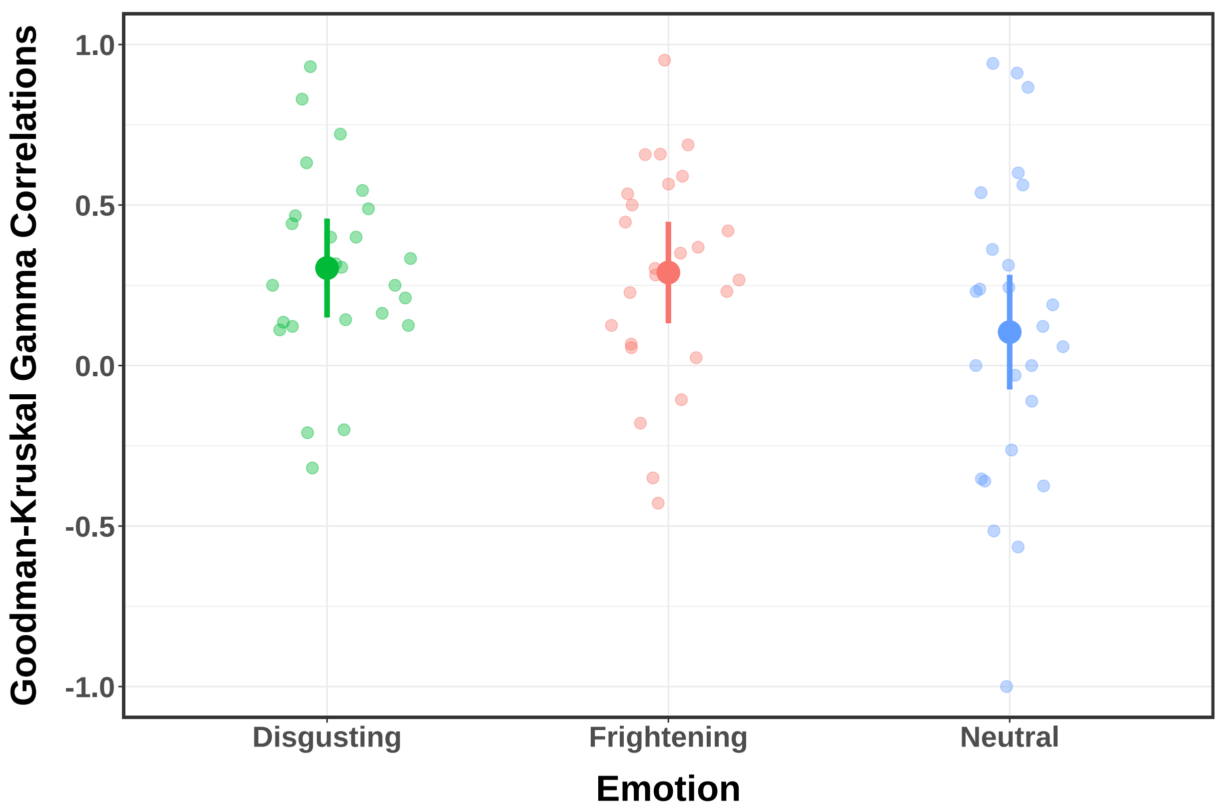
Resolution

Next, I analyzed *resolution*, a measure of relative metamemorial accuracy which quantifies the extent to which participants are able to discriminate between items which are subsequently remembered vs. forgotten. This measure is obtained by computing Goodman–Kruskal gamma correlations between JOLs and memory performance at the item-level (Nelson, 1984), with positive values indicating higher JOLs for remembered images, negative values indicating higher JOLs for forgotten images, and zero indicating chance-level metacognitive discrimination between remembered and forgotten images. I did not have any a priori hypotheses regarding resolution and analyzed this measure for completeness.

First, a repeated-measures ANOVA of resolution with Emotion (3) as a factor was conducted (see Figure 2). This analysis did not reveal a significant effect of Emotion ($F = 1.97$, $p = 0.151$), suggesting that participants' relative accuracy did not differ between emotions. Next, the resolution for each emotional category was compared to zero (i.e. chance-level discrimination) using a series of one-sample t-tests. Whereas the resolution associated with disgusting [$t(24) = 5.01$, $p < 0.001$] and frightening [$t(24) = 4.32$, $p < 0.001$] images was greater than zero, resolution for neutral images did not differ from zero [$t(24) = 1.07$, $p = 0.295$].

Figure 2

Resolution by Emotion in Experiment 1a



Note. Error bars represent 95% within-subject confidence intervals.

Stimulus Ratings

Finally, a series of repeated-measures ANOVAs with Emotion (3) as a factor were conducted on each of the four ratings to determine whether the images evoked the desired affective responses (see Table 1). A significant effect of Emotion on valence ratings [$F(2, 48) = 166.76$, $p < 0.001$, $\eta^2_p = 0.874$] indicated that neutral images were rated as more positively valenced than both frightening and disgusting images (both p's < 0.001). Frightening and disgusting images did not differ in valence ($p = 1.000$). There was a significant effect of Emotion on arousal ratings as well [$F(2, 48) = 180.15$, $p < 0.001$, $\eta^2_p = 0.882$]. Post-hoc tests revealed that neutral images were rated as less arousing than both frightening and disgusting images (both p's < 0.001). Notably, frightening images were rated as more arousing than disgusting images ($p < 0.001$). A significant effect of Emotion on fear ratings [$F(2, 48) = 183.49$, $p < 0.001$, $\eta^2 = 0.884$] revealed that neutral images were less frightening than frightening and disgusting images (both p's < 0.001), and frightening images were more frightening than disgusting images ($p < 0.001$). There was also a significant effect of Emotion on disgust ratings [$F(2, 48) = 201.57$, $p < 0.001$, $\eta^2_p = 0.894$]. Neutral images were rated as less disgusting than both frightening and disgusting images (both p's < 0.001). Additionally, disgusting images were rated as more disgusting than frightening images ($p < 0.001$). Lastly, a paired t-test comparing average fear ratings for frightening images to average disgust ratings for disgusting images revealed no significant difference in ratings [$t(24) = 0.406$, $p = 0.688$], suggesting that the frightening images were as frightening as the disgusting images were disgusting.

Table 1

Mean (SD) Valence, Arousal, Fear, and Disgust Ratings for Images in Experiment 1a

	Emotion Category		
Rating	Disgusting	Frightening	Neutral
Valence	2.25 (0.58)	2.11 (0.70)	4.60 (0.40)
Arousal	4.79 (0.61)	5.63 (0.62)	2.71 (0.72)
Fear	4.15 (1.28)	5.36 (0.90)	1.42 (0.48)
Disgust	5.30 (0.89)	3.04 (1.26)	1.24 (0.40)

Discussion

The results of Experiment 1a are consistent with some lines of research but not with others. In the present experiment, there was a significant effect of emotion on memory performance. This effect was characterized by the tendency for participants to remember both disgusting and frightening images better than neutral images. Such results are consistent with the emotional memory literature at large, which has consistently demonstrated that emotional information is better remembered than neutral information (for reviews, see Bennion et al., 2013; Buchanan, 2007).

The effect of emotion on JOLs observed in the present experiment is also consistent with prior research (e.g., Hourihan & Bursey, 2017; Hourihan et al., 2017; Nomi et al., 2013; Tauber & Dunlosky, 2012; Tauber et al., 2017; Witherby & Tauber, 2018; Zimmerman & Kelley, 2010). Much like the effect of emotion on actual memory performance, participants gave higher JOLs for emotional images relative to neutral images. Such results suggest that – consistent with the cue utilization framework (Koriat, 1997) – participants are indeed sensitive to the general cue of stimulus emotionality when predicting their memory performance and are (rightly) more confident in their ability to remember emotional images. Strikingly, the effect of emotion on JOLs was more than twice as large as the effect of emotion on actual memory (η^2_p = 0.718 vs. η^2_p = 0.317). Thus, it appears that participants might not simply be sensitive to emotion when making metamemorial judgments, they may be *oversensitive* to emotion.

Although these general effects of emotion on predicted and actual memory replicate prior research, the present experiment failed to replicate the disgust advantage in recall. Whereas prior research has suggested that disgusting information is better remembered than frightening information (Chapman, 2018; Chapman et al., 2013; Charash & McKay, 2002; Croucher et al., 2011; Ferré et al., 2018; Zhang et al., 2018), participants in the current study remembered disgusting and frightening images equally well. Possible reasons for this lack of replication are discussed in the following section (see Experiment 1b's Introduction).

As with actual memory, participants gave disgusting and frightening images equivalent JOLs. That is, although participants were more confident in their ability to remember both disgusting and frightening images relative to neutral images, participants did not differ in their confidence for disgusting and frightening images. Although it might therefore be tempting to conclude that participants are not sensitive to the cue of discrete emotion when predicting their memory performance, such a conclusion would be premature. As mentioned previously, in order to provide an appropriate test of whether participants are sensitive to discrete emotion when making JOLs, it would first be necessary to demonstrate that discrete emotions do indeed differ in their memorability. Indeed, the lack of a difference in memorability between discrete emotions in Witherby and Tauber's study (2018) is what motivated the present investigation. Without such a difference in memorability, it may be that participants are sensitive to discrete emotion, but only when discrete emotion actually affects memory. As such – due to the lack of a disgust advantage in the present experiment – I am unable to draw definitive conclusions about whether discrete emotions affect JOLs.

In addition to the primary results of this experiment, data on stimulus ratings and resolution were also analyzed. Regarding stimulus ratings, the results of the present experiment replicated Chapman's (2018; Experiment 1) results exactly: disgusting and frightening images did not differ in valence, frightening images were more frightening than disgusting images, disgusting images were more disgusting than frightening images, and frightening images were as frightening as disgusting images were disgusting. Such findings provide support for the stability of participants' appraisals of these images. Notably, the frightening images were rated as being significantly more arousing than the disgusting images. These results are consistent with Chapman's, who found that – although the two image categories were equated for arousal based on pilot data – the disgusting images were rated as less arousing than the frightening images during the memory experiment. The implications of this finding are discussed in the following section. As will be demonstrated in subsequent experiments, the consistently found difference between frightening and disgusting images in terms of arousal suggests that the images used by Chapman to represent these two emotional categories may not be properly equated in terms of arousal.

Lastly, analyses of participants' resolution demonstrated that although resolution was above chance for both frightening and disgusting images (but not neutral images), resolution did not differ as a function of emotion. This null result is consistent with the emotional metamemory literature at large, which has repeatedly found that emotion does not affect resolution (Hourihan & Bursey, 2017; Hourihan et al., 2017; Tauber & Dunlosky, 2012; Tauber et al., 2017; Witherby

& Tauber, 2018; Zimmerman & Kelley, 2010). As such, although emotion affects the magnitude of JOLs, it does not appear to influence participants' ability to discriminate between items which will and will not be subsequently remembered.

CHAPTER 3: EXPERIMENT 1B

The results of Experiment 1a failed to replicate the disgust advantage. As mentioned previously, this lack of replication precludes the investigation of discrete emotional effects on metamemory. As such, rather than investigating issues of metamemory as originally intended, the remainder of this thesis will be concerned with determining the replicability of the disgust advantage in memory.

There are a number of factors which might explain our failure to replicate the disgust advantage in Exp 1a. First, in Experiment 1a frightening images were rated as more arousing than disgusting images. Given that arousal has been found to increase memory independently of valence (e.g., Kensinger & Corkin, 2004), one might wonder whether the increased arousal of frightening images might have prevented the disgust advantage from manifesting. Although plausible, given that Chapman (2018; Experiment 1) observed the disgust advantage using the same materials used here in a sample which also rated the frightening images as more arousing than the disgusting images, this explanation seems unlikely. That is, because Chapman's results demonstrate that the disgust advantage can still be found when using frightening and disgusting images which are not equated for arousal, this confound is likely not responsible for our failure to replicate this effect.[6]

Another, more likely explanation concerns the effects of making JOLs on memory performance. Prior research has found that asking participants to make JOLs sometimes

[6] The results of Experiment 2 also argue against this explanation in that the disgust advantage was observed despite a significant difference between frightening and disgusting images in terms of arousal.

eliminates otherwise robust memory effects. For example, although information which has been perceptually degraded is typically better remembered than perceptually intact information (e.g., Nairne, 1988), Besken and Mulligan (2013) found that having participants make JOLs after each item eliminated this effect. Similarly, JOLs have been found to eliminate the generation effect (Begg et al., 1991; Matvey et al., 2001), an effect whereby information which participants generate themselves is better remembered than information which is not self-generated (Slamecka & Graf, 1978).

Although Experiment 1b was not designed to test the exact mechanisms by which JOLs may have eliminated the disgust advantage, prior research suggests two possibilities. It has been suggested that JOLs might eliminate memory effects either because 1) JOLs encourage deep processing of the less memorable class of materials, thereby equalizing memory performance and eliminating differences between categories of materials, or 2) JOLs divert attention away from the studied materials and towards the JOLs (Besken & Mulligan, 2013). Within the context of the present investigation, JOLs might have eliminated the disgust advantage by increasing the depth of processing of frightening images, or by diverting attention away from the disgusting images. Rather than investigating whether these specific explanations apply to the current findings, the goal of Experiment 1b was to first establish whether JOLs do indeed eliminate the disgust advantage.

Given these considerations, the purpose of Experiment 1b was to test whether the failure to replicate the disgust advantage in Experiment 1a was due to the presence of JOLs. As such, the design of Experiment 1b was similar to that of Experiment 1a except that participants were no longer asked to make JOLs. This experiment therefore addressed questions both about the reactivity of JOLs, and about the replicability and potential boundary conditions for the disgust advantage.

Method

Participants[7]

As in Experiment 1a, participants were 25 undergraduate students who were compensated for their participation with course credit.

Design, Materials, and Procedure

The methods were identical to Experiment 1a with the following exception: participants in Experiment 1b no longer made JOLs during encoding. As such, during the encoding phase participants simply studied each image for 2 s in preparation for an upcoming memory test. As in Experiment 1a, encoding was intentional.

Results

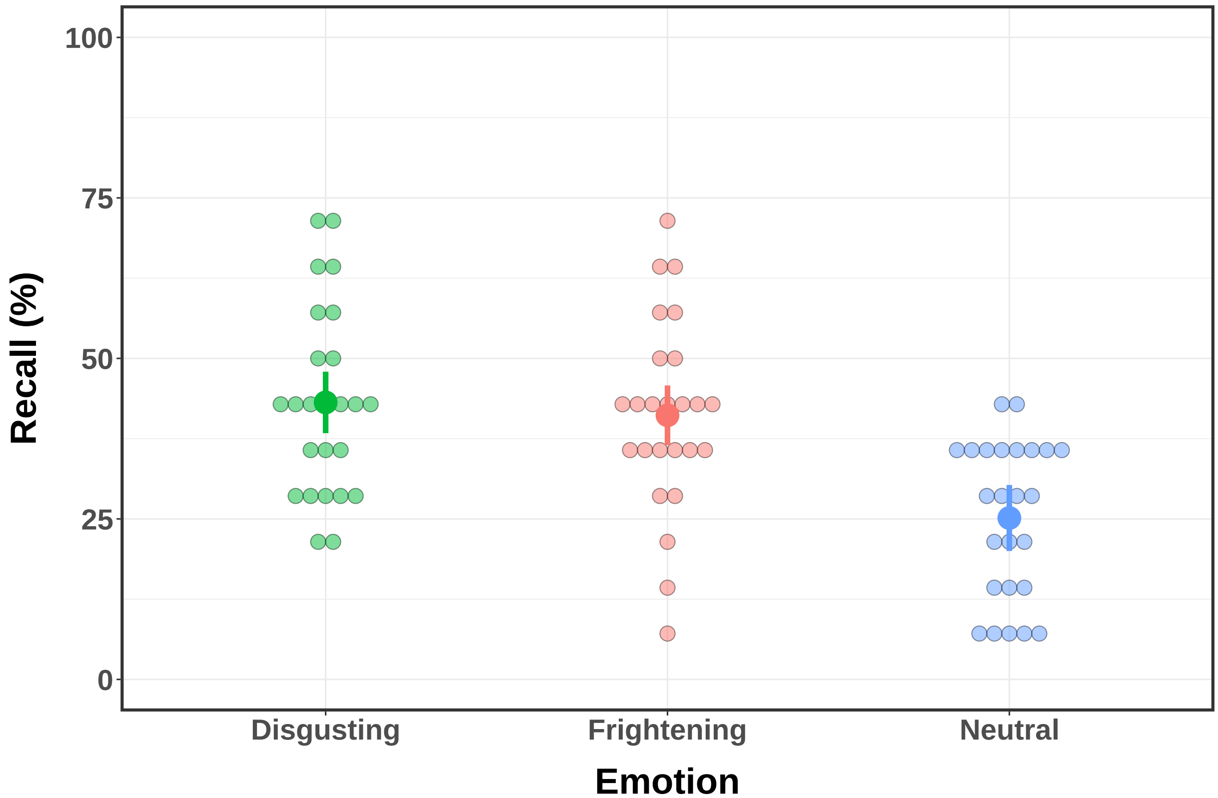
Recall

Recall was analyzed using a repeated-measures ANOVA with Emotion (3) as a factor (see Figure 3). The results of Experiment 1b replicate those of Experiment 1a. Specifically, there was a significant effect of Emotion on recall performance [$F(2, 48) = 17.53$, $p < 0.001$, $\eta^2_p = 0.422$]. Post-hoc revealed that participants recalled fewer neutral images compared to both disgusting and frightening images (both p's < 0.001). Participants did not differ in their recall of disgusting and frightening images ($p = 1.000$).

[7] One participant in Experiment 1b became distressed upon seeing the emotional images during the encoding phase and terminated the experiment. This subject was replaced.

Figure 3

Recall by Emotion in Experiment 1b



Note. Error bars represent 95% within-subject confidence intervals.

Stimulus Ratings

Next, stimulus ratings for valence, arousal, fear, and disgust were analyzed using a series of repeated-measures ANOVAs with Emotion (3) as a factor (see Table 2). As with recall, the stimulus ratings replicated the results of Experiment 1a exactly. Specifically, there was a significant effect of Emotion on valence ratings [$F(1.58, 37.88) = 154.81$, $p < 0.001$, $\eta^2_p = 0.866$], indicating that neutral images were rated more positively than both frightening and disgusting images (both p's < 0.001). Frightening and disgusting images did not differ in their valence ($p = 1.000$). There was also a significant effect of Emotion on arousal ratings [$F(2, 48) = 139.74$, $p < 0.001$, $\eta^2_p = 0.853$]. Post-hoc tests revealed that neutral images were rated as less arousing than both frightening and disgusting images (both p's < 0.001). Frightening images were rated as more arousing than disgusting images ($p < 0.001$). There was a significant effect of Emotion on fear ratings [$F(2, 48) = 181.81$, $p < 0.001$, $\eta^2_p = 0.883$]. Frightening images were rated as more frightening than disgusting images ($p < 0.001$), and neutral images were rated as less frightening than both frightening and disgusting images (both p's < 0.001). Similarly, there was a significant effect of Emotion on disgust ratings [$F(2, 48) = 172.94$, $p < 0.001$, $\eta^2_p = 0.878$]. Disgusting images were rated as more disgusting than frightening images ($p < 0.001$), and neutral images were rated as less disgusting than both frightening and disgusting images (both p's < 0.001). Lastly, a paired t-test indicated that the frightening images were as frightening as the disgusting images were disgusting [$t(24) = -0.217$, $p = 0.830$].

Table 2

Mean (SD) Valence, Arousal, Fear, and Disgust Ratings for Images in Experiment 1b

	Emotion Category		
Rating	Disgusting	Frightening	Neutral
Valence	2.41 (0.51)	2.31 (0.56)	4.89 (0.65)
Arousal	4.95 (0.63)	5.68 (0.61)	2.72 (0.89)
Fear	3.89 (1.32)	5.12 (1.08)	1.29 (0.38)
Disgust	5.16 (0.98)	3.07 (1.28)	1.11 (0.20)

Discussion

The goal of Experiment 1b was to investigate whether the failure to replicate the disgust advantage in Experiment 1a was due to the inclusion of JOLs. This hypothesis was motivated by research demonstrating that item-by-item JOLs sometimes eliminate well-established memory effects (e.g., Begg et al., 1991; Besken & Mulligan, 2013; Matvey et al., 2001). However, the results of Experiment 1b provide evidence against this hypothesis; despite producing a general mnemonic advantage for emotional relative to neutral images, the results again failed to replicate the disgust advantage.

CHAPTER 4: EXPERIMENT 2

The results of the prior two experiments suggest that the disgust advantage may either not be replicable, or that the disgust advantage may not be robust to seemingly minor changes in experimental procedures. However, before drawing such a conclusion, it is important that remaining explanations for the current failure to replicate be ruled out. In particular, Experiment 1b differed from Chapman's (2018; Experiment 1) recent demonstration of the disgust advantage in two ways: 1) in Experiment 1b encoding was intentional, whereas in Chapman's experiment encoding was incidental, and 2) in Experiment 1b participants simply studied the images, whereas in Chapman's experiment participants completed a *line discrimination task* (LDT) meant to assess attention bias toward the emotional images. The potential importance of these differences is considered below.

In Chapman's study (2018; Experiment 1) participants encoded the images incidentally, without any awareness of the impending recall test. In contrast, in both Experiments 1a and 1b participants engaged in intentional encoding; participants were aware that their memory for the images would be tested later. This, of course, was necessary in order to have subjects provide JOLs and assess their metamemory. Prior to conducting the present experiment, it was unclear whether incidental encoding was necessary to observe the disgust advantage, as nearly every prior demonstration of the disgust advantage has used incidental encoding (Chapman, 2018; Chapman et al., 2013; Charash & McKay, 2002; Croucher et al., 2011; Ferré et al., 2018, but see Zhang et al., 2018), and no study to my knowledge has investigated the effects of varying encoding instructions on the disgust advantage directly. The present experiment is therefore the first to systematically investigate the effects of encoding instructions on the disgust advantage. Because intentional encoding is thought to encourage effortful, elaborative processing to a greater extent than incidental encoding (e.g., Block, 2009; Naveh-Benjamin et al., 2014; Talmi, 2013), it stands to reason that if the disgust advantage is mediated by such controlled processing, then intentional encoding may eliminate this effect by encouraging the elaboration of both disgusting and frightening images equally.

In addition to the encoding instructions, Experiments 1a and 1b differed from Chapman's (2018; Experiment 1) experiment in that participants in Chapman's study completed the LDT during encoding. Briefly, the LDT entails presenting horizontal lines either above or below the studied images. Participants are tasked with indicating the location of the line as quickly as possible (see Method below). To the extent that emotional images capture attention, one would expect these images to interfere with task performance, resulting in slower reaction times for emotional compared to neutral trials. Because a prior study which manipulated whether or not participants completed the LDT during encoding found that the disgust advantage appeared in both groups (Chapman et al., 2013; Experiment 3), it was not predicted that this methodological detail would be necessary to observe the effect. Even so, the present experiment included the LDT to more closely replicate Chapman's (2018; Experiment 1) procedures.

A primary goal of Experiment 2 was therefore to serve as a direct replication of Chapman's (2018) Experiment 1, the experiment on which the present study was based, in order to provide a stronger test of whether the disgust advantage is replicable. In Experiment 2, encoding (intention vs. incidental) was manipulated between-subjects such that one group was made aware of the upcoming memory test and the other group was unaware of the memory test.

Both groups completed the LDT. The intentional condition provides a direct replication of Chapman's procedures, and the addition of an intentional encoding group allows me to isolate the cause of a successful replication, should one occur. If the study included only an incidental group who completing the LDT (the direct replication of Chapman's procedures) and subsequently replicated the disgust advantage, it would be unclear whether incidental encoding or the LDT were responsible for the successful replication.

The current design allows for more informative conclusions. If the disgust advantage should occur for the incidental group but not the intentional group, this would suggest that incidental encoding is necessary to observe the disgust advantage. If the effect should occur for both groups, this would suggest that the LDT is necessary, as Experiment 1b was identical to the intentional condition except for the LDT. In contrast, if the disgust advantage does not occur for either group, this would suggest that the disgust advantage is not replicable even in a quite narrow sense, as the incidental group is effectively a direct replication of Chapman's (2018; Experiment 1) procedures.

In addition to assessing the mnemonic effects of discrete emotions, the inclusion of the LDT in Experiment 2 allowed me to investigate the attentional effects of discrete emotions. Chapman (2018; Experiment 1) found evidence using the LDT that participants allocate greater attention to disgusting images compared to frightening images, and suggested that these attentional differences mediate the mnemonic advantage for disgust (see Chapman et al., 2013 for similar results using the LDT). The claim that disgusting information attracts greater attention than frightening information is consistent with past research (e.g., Carretié et al., 2011; Ciesielski et al., 2010; Cisler et al., 2009; Krusemark & Li, 2011; Van Hooff et al., 2013, but see Zimmer et al., 2016; Zhang et al., 2017). As such, a secondary goal of Experiment 2 was to test the replicability of Chapman's (2018; Experiment 1) LDT results, and build upon research on disgust and attention more broadly.

Lastly, the manipulation of encoding allowed for the examination of an additional issue of interest: the effects of encoding instructions on emotional memory in general. It has been claimed that emotionally induced memory effects are stronger during incidental compared to intentional encoding, as intentional encoding may encourage the deep, elaborative processing of non-emotional materials and – as a result – increase their memorability to a level on par with emotional materials (Talmi, 2013). Indeed, such a claim is supported by research demonstrating that the mnemonic advantage for emotional relative to neutral information is larger under conditions of shallow encoding (Jay et al., 2008; Ritchey et al., 2011), which incidental encoding is thought to encourage.

Yet despite the intuitive nature of this argument, what little evidence exists which directly speaks to this claim is somewhat mixed. Although D'Argembeau and Van der Linden (2004) found that memory for contextual information was better for emotional relative to neutral words during incidental but not intention encoding[8], the effect of emotion on spatial memory was not moderated by encoding instructions. In contrast, Preuß et al. (2009) manipulated encoding instructions (intentional vs. incidental) between-subjects and found that the effect of emotion on recall performance (i.e. item memory) did not interact with encoding instructions, suggesting that the size of the mnemonic advantage of emotion did not differ between conditions. Thus, given the claim that emotionally enhanced memory is stronger during incidental encoding and the paucity of evidence which directly supports this claim, a secondary goal of Experiment 2 was to test the effect of encoding instructions on emotionally enhanced memory.

---

[8] In D'Argembeau and Van der Linden's (2004) studies, encoding instructions manipulated whether participants attempted to memorize the color of the word, rather than the word itself.

Method[9]

Participants

In the prior two experiments, we determined our sample size using a priori power analyses based on the main effect of emotion on recall reported by Chapman (2018; Experiment; $\eta^2_p = 0.33$). Thus far we have consistently observed this main effect of emotion. For Experiment 2, we chose instead to determine our sample size based on the effect size of the difference between recall for disgusting and frightening images, as this difference was the primary effect of interest. Based on this effect size ($d_z = 0.51$), to achieve at least 80% power requires 33 subjects in each group. As such, I initially planned to collect 35 subject per group. However, because the COVID-19 pandemic forced me to stop collecting data midway through the semester, the present experiment contains data from 62 undergraduate students who were compensated for their participation with course credit (31 per group). Post-hoc power analyses indicated that this sample size resulted in 78.46% power in each group to detect Chapman's difference in recall between disgusting and frightening images.

Design and Materials

The design of Experiment 2 was a 3 (Emotion: disgust, fear, neutral; within-subjects) x 2 (Encoding: intentional vs. incidental; between-subjects) mixed design. Studied images were identical to prior experiments.

Procedure

Participants began by completing the encoding phase. Participants were told that during this task, they would see a series of images appear on the screen, each with a horizontal line either above it or below it. Participants were instructed to indicate the location of this line as

[9] Preregistration materials for Experiment 2 are available at: https://aspredicted.org/blind.php?x=n3pz4y

quickly and accurately as possible by pressing the up arrow if the line appeared above the image, and the down arrow if the line appeared below the image. Additionally, participants were asked to keep their eyes focused on the images throughout the duration of the task.

As in prior experiments, images were displayed for 2 s and were separated by a 500 ms interstimulus interval. Each image was displayed for 2 s regardless of how quickly participants made the response to the LDT, and participants did not receive feedback on their performance. The dimensions of the horizontal lines in the present study were chosen to match the dimensions of Chapman's (2018; Experiment 1) lines relative to the size of the images in our prior two experiments. More specifically, the proportionate size of the lines relative to the images was $0.375 \times 0.008\overline{3}$. Each line orientation (above vs. below) appeared equally often within each of the three emotional categories. Additionally, line position was counterbalanced such that a given image had a horizontal line above it in one list and below it in the other list. As in prior experiments, images appeared in a new random order for each participant.

For participants in the intentional group, the experimenter explained that – in addition to completing the LDT – participants were to study the images for a memory test which would occur 45 minutes after the LDT. In contrast, participants in the incidental group were not informed that there would be a memory test and were simply instructed to complete the LDT.

The remainder of the experimental procedures (i.e. the retention interval, the recall test, and the stimulus rating phase) were identical to prior experiments.

Data Analysis

Reaction time data for the LDT were trimmed according to the procedures described by Chapman (2018; Experiment 1). Specifically, only reaction times for correct trials were analyzed. Removing trials with incorrect responses resulted in a loss of 3.02% of trials (not including the two excluded participants, see Footnote 10). Additionally, reaction times which were more than two standard deviations away from each participant's grand mean were excluded. Removing these extreme trials resulted in a loss of an additional 4.26% of trials. In total, 7.14% of trials were removed prior to conducting the reaction time analyses. Similar to Chapman, mean reaction times were used in the LDT analyses.

Results

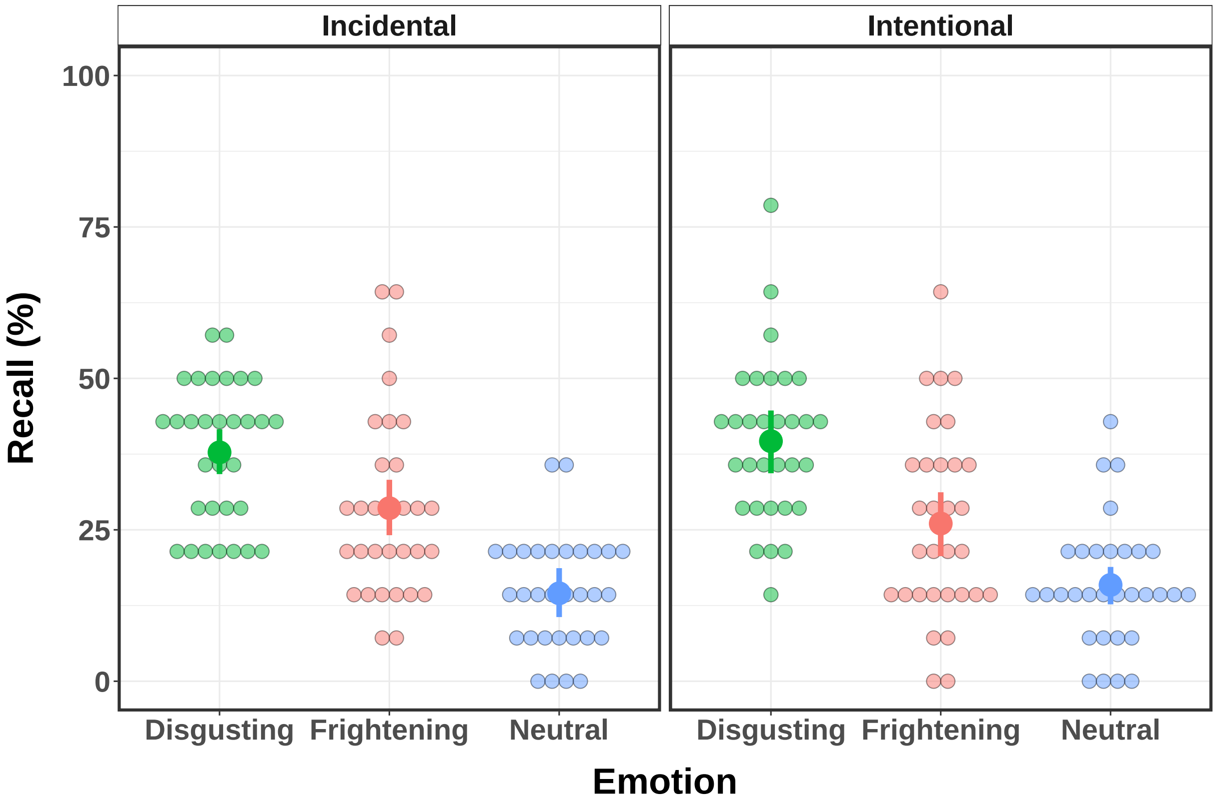
Recall[10]

I began by analyzing the proportion of pictures recalled with a 3 (Emotion: disgusting, frightening, neutral; within-subjects) x 2 (Encoding: intentional vs. incidental; between-subjects) mixed ANOVA (see Figure 4). There was a significant effect of Emotion [$F(1.83, 109.99) = 60.05$, $p < 0.001$, $\eta^2_p = 0.500$]. As in previous experiments, both frightening and disgusting images were better remembered than neutral images (both p's < 0.001). In this experiment – however – the disgust advantage was replicated: disgusting images were better remembered than frightening images ($p < 0.001$). The effects of Encoding and the Encoding by Emotion interaction were both non-significant ($F = 0.01$, $p = 0.915$ and $F = 0.63$, $p = 0.522$, respectively).

[10] One participant who became uncomfortable upon seeing the emotional images requested to take a break midway through the encoding phase. Due to an error in the experimental program, the program terminated instead of resuming where this participant left off in the LDT, and this participant completed the LDT in its entirety from the beginning. As such, this participant received additional exposure to some of the studied images. Because sensitivity analyses indicated that the results of this experiment did not change when this participant was excluded, this participant's data was retained in the analysis of recall data.

Figure 4

Recall by Emotion and Encoding in Experiment 2



Note. Error bars represent 95% within-subject confidence intervals.

Line Discrimination Task[11]

Reaction times were submitted to a 3 (Emotion: disgusting, frightening, neutral; within-subjects) x 2 (Encoding: intentional vs. incidental; between-subjects) mixed ANOVA (see Figure 5). There was a marginal effect of Encoding, indicating that participants were slower in the intentional condition compared to the incidental condition [$F(1, 58) = 3.98$, $p = 0.051$, η^2_p =

[11] One participant had an accuracy of 0% because this participant never made a response. Another participant with an accuracy of 0% appears to have reversed the keys associated with the corresponding correct response (e.g., the participant pressed the up arrow for trials in which the line appeared below the image). As such, both subjects were excluded from all LDT analyses. Because excluding these participants did not change the recall results, these participants were included in the previously-reported recall analyses.

0.064]. There was also a main effect of Emotion [$F(2, 116) = 15.76$, $p < 0.001$, $\eta^2_p = 0.214$]. Post-hoc tests showed that participants were slower to respond during trials containing both frightening and disgusting images compared to trials containing neutral images (both p's < 0.001). However, participants did not differ in their reaction times for trials containing frightening and disgusting images ($p = 1.000$). Additionally, the Encoding by Emotion interaction was non-significant ($F = 0.47$, $p = 0.629$).

Table 3

Mean (SD) Reaction Times (ms) for the LDT by Emotion and Encoding in Experiment 2

	Emotion Category		
Encoding	Disgusting	Frightening	Neutral
Incidental	714.84 (269.49)	701.24 (286.10)	669.04 (264.82)
Intentional	865.41 (325.76)	863.02 (326.98)	812.88 (309.55)

Stimulus Ratings[12]

Next, the stimulus ratings for valence, arousal, fear, and disgust were analyzed using repeated-measures ANOVAs with Emotion (3) as a factor (see Table 3). There was a significant effect of Emotion on valence ratings [$F(1.79, 104.06) = 327.91$, $p < 0.001$, $\eta^2_p = 0.850$]. Post-hoc tests revealed that neutral images were rated as more positively valenced than both frightening and disgusting images (both p's < 0.001). Frightening and disgusting images did not differ in valence ($p = 1.000$). Similarly, there was a significant effect of Emotion on arousal ratings [$F(1.70, 98.68) = 205.15$, $p < 0.001$, $\eta^2_p = 0.780$]. Post-hoc tests revealed that neutral images

[12] One participant who reported frequently reversing the rating scale for valence was excluded from analyses of stimulus ratings. Additionally, two subjects became distressed and left prior to the stimulus rating phase (but after the recall phase).

were rated as less arousing than both frightening and disgusting images (both *p*'s < 0.001). Additionally, frightening images were rated as more arousing than disgusting images ($p < 0.001$). There was a significant effect of Emotion on fear ratings [$F(2, 116) = 351.43$, $p < 001$, η^2_p = 0.858]. Neutral images were rated as being less frightening than both frightening and disgusting images (both *p*'s < 0.001), and frightening images were rated as more frightening than disgusting images ($p < 0.001$). Lastly, there was a significant effect of Emotion on disgust ratings [$F(2, 116) = 334.75$, $p < 0.001$, $\eta^2_p = 0.852$]. Neutral images were rated as less disgusting than both frightening and disgusting images (both *p*'s < 0.001), and disgusting images were rated as more disgusting than frightening images ($p < 0.001$). Finally, a paired t-test found that – unlike in Experiments 1a and 1b – disgusting images were rated as more disgusting than frightening images were rated as frightening [$t(58) = 2.15$, $p = 0.036$].

Table 4

Mean (SD) Valence, Arousal, Fear, and Disgust Ratings for Images in Experiment 2

	Emotion Category		
Rating	Disgusting	Frightening	Neutral
Valence	2.25 (0.64)	2.28 (0.60)	4.72 (0.57)
Arousal	4.58 (0.80)	5.29 (0.77)	2.76 (0.70)
Fear	4.02 (1.25)	5.00 (1.13)	1.38 (0.46)
Disgust	5.28 (1.08)	3.02 (1.30)	1.24 (0.39)

Discussion

The primary result of Experiment 2 was that – unlike in Experiments 1a and 1b – the disgust advantage was successfully replicated. This effect was characterized by greater recall performance for disgusting compared to frightening images. Because the disgust advantage appeared regardless of whether participants engaged in intentional or incidental encoding (i.e. the effect of emotion did not interact with encoding type), it appears that incidental encoding is not necessary to replicate this effect. Instead, it appears that the reason that the disgust advantage was not replicated in prior experiments was that these experiments did not utilize the LDT at encoding. I will refrain from speculating as to why this methodological detail is of importance until the General Discussion. Although the disgust advantage appears to be a replicable phenomenon, the present investigation suggests that boundary conditions such the task which participants engage in during encoding may constrain its appearance.

The results of the LDT revealed that participants were slower to indicate the location of the line during trials with emotional images compared to trials with neutral images. Such results suggest that the emotional images in the present study captured participants' attention to a greater extent than the neutral images did. These results are consistent with prior research, which has consistently demonstrated that emotional information captures attention and interferes with ongoing task performance to a greater extent than non-emotional information (for a review, see Carretié, 2014).

Although the results of the LDT are consistent with the literature on emotion and attention at large, Experiment 2's LDT results are inconsistent with those reported by Chapman. Specifically, although Chapman has found that participants are slower during trials which contain disgusting images compared to trials which contain frightening images (e.g., Chapman, 2018; Chapman et al., 2013), the present experiment found no significant difference between these emotional categories in terms of task interference. Broadly speaking, although some prior research indicates that disgusting information captures more attention than frightening information (Carretié et al., 2011; Ciesielski et al., 2010; Cisler et al., 2009; Krusemark & Li, 2011; Van Hooff et al., 2013), this is not always found (Zimmer et al., 2016; Zhang et al., 2017). For example, in Zhang et al.'s study participants viewed briefly presented pairs of emotional and neutral faces and responded to geometric targets presented behind one face or the other. Whereas reaction times for trials with disgusted faces were quicker if the object was positioned on the side of the screen which the neutral face had previously occupied, the opposite was found for trials with fearful faces. Such results suggest that – at least under some circumstances – frightening information may capture attention to a greater extent than disgusting information (see Zimmer et al., 2016 for similar results using auditory stimuli). Reconciling these disparate results will be an important challenge for researchers in this domain.

Alternatively, it is possible that the lack of a difference in attention capture by disgusting and frightening images in the present study is due to the greater arousal of the frightening images used. That is, although disgusting information may be more attention capturing than frightening information, the competing effects of arousal on attention may have obscured my ability to detect this effect. Support for this conclusion comes from a study by Schimmack and Derryberry (2005), who found that negative images with higher arousal slow line detection to a greater extent than negative images with lower arousal. As such, it appears that participants' attention is sensitive to differences in stimulus arousal, and it is therefore possible that the confound in arousal between disgusting and frightening images explains our inability to replicate Chapman's LDT results (Chapman, 2018; Chapman et al., 2013). Even so, Chapman (2018; Experiment 1) found greater task interference for disgusting images relative to frightening images using the LDT regardless of the fact that participants rated the frightening images as more arousing. Although it is therefore plausible that the confound in arousal between disgusting and frightening images interfered with our ability to replicate Chapman's LDT results, prior research indicates that this explanation may not fully account for the failure to replicate.

Whatever the reason for our failure to replicate the increased attention bias toward disgusting images, it is interesting to note that the mnemonic advantage for disgust was observed in the absence of a significant difference in attentional bias between frightening and disgusting images during encoding. Such results suggest that greater attention for disgusting relative to frightening images is not a prerequisite for observing the mnemonic advantage for disgust. If this is the case, it may be that attentional bias is one of multiple mediators of the disgust advantage, with differential attentional bias towards frightening and disgusting information contributing to and perhaps amplifying the disgust advantage when present (Chapman, 2018; Experiment 1), but not preventing the disgust advantage when absent.

A final goal of Experiment 2 was to test the claim that emotionally enhanced memory effects are stronger under incidental compared to intentional encoding (Talmi, 2013). This claim seems convincing given the argument that incidental encoding may encourage shallow processing, coupled with research demonstrating that emotionally enhanced memory effects are larger when participants engage in shallow encoding (Jay et al., 2008; Ritchey et al., 2011). Even so, the results of Experiment 2 argue against this claim, as the size of the mnemonic advantage for emotion did not differ between encoding conditions as demonstrated by the nonsignificant interaction between emotion and encoding type. Such results mirror the findings of Preuß et al. (2009), who similarly manipulated encoding type between-subjects and failed to observe an interaction between emotion and encoding type on participants' recall performance for images. Although it is possible that incidental encoding enhances the effects of emotion on memory in certain circumstances or for certain types of materials, the present results cast doubt on the generality of this claim.

CHAPTER 5: GENERAL DISCUSSION

Initially, the goal of the present investigation was to investigate the effects of stimulus emotionality on metamemorial judgments. More specifically, I sought to determine whether participants are sensitive to the mnemonic effects of discrete emotions when predicting their memory performance. This question was motivated by research demonstrating that participants are sensitive to mnemonic cues that reflect aspects of the stimulus to be remembered (Koriat, 1997), including whether the stimulus is emotional or non-emotional (Hourihan & Bursey, 2017; Hourihan et al., 2017; Nomi et al., 2013; Tauber & Dunlosky, 2012; Tauber et al., 2017; Witherby & Tauber, 2018; Zimmerman & Kelley, 2010). Given that discrete emotions have been shown to affect a participant's likelihood of remembering a stimulus above and beyond affective dimensions such as valence and arousal with disgusting stimuli being better remembered than frightening stimuli (Chapman, 2018; Chapman et al., 2013; Charash & McKay, 2002; Croucher et al., 2011; Ferré et al., 2018; Zhang et al., 2018), the present investigation sought to determine whether participants would incorporate this mnemonic advantage for disgusting information into their metamemorial judgments. However, because the disgust advantage did not replicate in Experiment 1a, subsequent experiments were concerned with determining the reasons for this failure to replicate.

Experiment 1b was designed to test whether the initial failure to replicate the disgust advantage was due to the inclusion of JOLs. It was reasoned that – because the presence of JOLs has been shown to eliminate otherwise robust memory effects (e.g., Begg et al., 1991; Besken &

Mulligan, 2013; Matvey et al., 2001) – the inclusion of JOLs in Experiment 1a may have interfered with the appearance of the disgust advantage. However, this hypothesis was proven unlikely, as Experiment 1a likewise failed to replicate the disgust advantage despite the removal of JOLs.

Experiment 2 was therefore designed to serve as a more direct replication of Chapman's (2018; Experiment 1) procedures; in Experiment 2, participants completed the LDT meant to assess attention biases toward the emotional images. Additionally, encoding type was manipulated such that one group was made aware of the impending memory test and another was unaware of the memory test. Because the disgust advantage was replicated regardless of whether encoding was intentional or incidental, the results of Experiment 2 suggest that the LDT may be necessary to observe the disgust advantage using the present materials. In order to provide additional support for the notion that the LDT is necessary to observing the disgust advantage, I compared memory for disgusting vs. frightening images across experiments using a 2 (Experiment: 2 vs. 1b) by 2 (Emotion: disgusting vs. frightening) mixed ANOVA. Consistent with the idea that the LDT contributed to the detection of the disgust advantage, there was a significant interaction between Experiment and Emotion [$F(1, 85) = 4.64$, $p = 0.034$, $\eta^2_p = 0.052$], indicating that whereas participants remembered more disgusting compared to frightening images in Experiment 2 [$F(1, 61) = 21.39$, $p < 0.001$, $\eta^2_p = 0.260$], recall did not differ between the two categories in Experiment 1b ($F = 0.41$, $p = 0.529$).

Because Chapman et al. (2013; Experiment 3) found the disgust advantage for recognition memory regardless of whether participants completed the LDT, it was not predicted a priori that this methodological detail was crucial to observe the effect. As such, the speculations that follow as to why this detail may have accounted for the initial failures to replicate the effect are post hoc.

The most plausible explanation as to why the LDT may be necessary to observe the disgust advantage has to do with the nature of the LDT. In essence, the LDT might be thought of as a form of *divided attention* (DA), in which participants must simultaneously focus on each centrally presented image and at the same time indicate the location of the line. As such, participants who encode images while completing the LDT experience a situation in which their attentional resources are divided between the ongoing LDT and the images onscreen. It may be that this type of encoding impairs memory for frightening information to a greater degree than for disgusting information. In support of this claim, the majority of studies which have demonstrated the disgust advantage have utilized some form of DA task at encoding (e.g., Chapman, 2018; Chapman et al., 2013; Charash & McKay, 2002; Ferré et al., 2018, but see Croucher et al., 2011; Zhang et al., 2018).

The claim that the LDT – and by extension DA – is necessary to observe the disgust advantage seems at odds with Chapman et al.'s (2013; Experiment 3) results in that these authors found the disgust advantage regardless of whether participants engaged in the LDT. One potential explanation for these discrepant results is that the disgust advantage simply may not be replicable under conditions of full attention. Future work addressing this possibility seems warranted. Another potential explanation relates to a difference between Chapman et al.'s procedures and those of the present investigation: participants in Chapman et al.'s study completed a recognition test, whereas participants in the present studies completed a recall test. Prior research has shown that recognition memory is less resource demanding than recall, and – at least when attention is manipulated at retrieval – recognition tends to be less sensitive to the effects of DA on memory performance (Baddeley et al., 1984; Craik et al., 1996). Because recognition tests are less sensitive to the effects of DA than recall tests, it is possible that whether or not participants encode information under full or divided attention is more important when examining mnemonic effects such as the disgust advantage using recall tests compared to recognition tests.

Why might it be that the disgust advantage is more likely to manifest when participants encode information under situations of DA? First, if – as suggested by prior research (Carretié et al., 2011; Ciesielski et al., 2010; Cisler et al., 2009; Krusemark & Li, 2011; Van Hooff et al., 2013) – disgusting information truly attracts attention to a greater degree than frightening information, it stands to reason that distracting secondary tasks may be less likely to pull participants' attention away from disgusting information. In other words, attending to disgusting information may be somewhat compulsory, with the attentional system giving high priority to disgust even when participants are confronted with effortful secondary tasks. If the increased attentional capture of disgusting information relative to frightening information makes disgusting information more resistant to the harmful effects of DA at encoding, one would expect the disgust advantage in memory to be stronger in situations of DA.

Although this explanation is plausible, it is difficult to reconcile with the lack of a significant difference in attention toward frightening and disgusting images in Experiment 2. Alternatively, it may be that the mechanisms which mediate disgust-memory are different from those that mediate fear-memory, and that the mechanisms mediating the former are less dependent on the availability of attentional resources than those mediating the latter. Stated differently, the mechanisms which underly disgust-memory may be less controlled and resource-demanding than those which underly fear-memory. In support of the idea that memory for different types of emotional material differs in terms of automaticity, Kensinger and Corkin (2004) found that memory for negative non-arousing material is disrupted by DA whereas memory for negative arousing material is not. It may therefore be that the mechanisms which mediate disgust-memory are relatively more resistant to DA than those mediating fear-memory.[13]

If disgust-memory does indeed rely upon processes which are relatively less resource-dependent than fear-memory, it is natural to wonder what these mechanisms might be. Although a number of authors have suggested that the disgust advantage may be due to increased attention at encoding for disgusting compared to frightening information (Chapman, 2018; Marchewka et al., 2016, but see Chapman et al., 2013), other candidate mechanisms have been put forth. For example, some have suggested that disgusting information may be more memorable than frightening information because it is higher in *impact* (Chapman et al., 2013; Croucher et al., 2011), a construct originating from photojournalism which characterizes images which have a striking, eye-catching nature, perhaps because these images are appraised as being of immediate significance to the observer (Ewbank et al., 2009; Murphy et al., 2010). Given the paucity of research on emotional memory and impact (e.g., Ramponi et al., 2010), an important challenge for future research will be to better characterize the role of impact as a mediator of the disgust advantage, as well as to determine whether the effects of impact on disgust-memory operate in a relatively automatic, resource-sparing manner. Alternatively, given that images high in impact have been defined as those which are "eye-catching" (Murphy et al., 2010), it is possible that impact simply measures participants' subjective awareness that certain emotionally charged

[13] The fact that more arousing stimuli were more resistant to DA is somewhat problematic for the claim that differential resistance to DA at encoding between disgusting and frightening information explains our prior failures to replicate the disgust advantage given that the frightening images in the present study were consistently rated as more arousing than the disgusting images. However, it is possible that whatever aspects of disgusting stimuli cause them to be particularly DA-resistant were sufficiently strong as to overcome the confound of arousal in the present study.

images are likely to capture their attention. Future research should test this possibility by investigating whether impact mediates the disgust advantage or has effects on emotional memory more generally after taking objective measures of attention bias into account.

When considering the differences between Experiment 2 and Experiments 1a and 1b, one final difference of potential importance remains: Experiment 2 was conducted during the outbreak and escalation of the COVID-19 pandemic. More specifically, data collection for Experiment 2 occurred between January 15, 2020 and March 6, 2020. Although speculative, it is plausible that during times of pandemic, people are more sensitive to disgust. The amount of disgust evoked by an image of a medical procedure or a contaminant – for example – may be elevated during periods of time in which threats of illness are particularly salient in the general public. If this is the case, it is possible that such heightened disgust sensitivity may have downstream consequences on the processing of disgusting stimuli, perhaps increasing their memorability in a way which facilitated the detection of the disgust advantage in Experiment 2. To test this hypothesis, I conducted an analysis based on the date that participants completed the study, in which participants who participated between January 15 and February 7 were assigned to an "early" group and participants who participated between February 10 and March 6 were assigned to a "late" group. A mixed ANOVA of the recall data from Experiment 2 with Emotion (3; within-subjects) and Date (2: early vs. late; between-subjects) revealed no effect of Date (F = 0.24, p = 0.628) and a nonsignificant interaction between Date and Emotion (F = 0.11, p = 0.879). Additionally, I analyzed the average disgust recall data using a regression analysis with Date as a continuous (i.e. non-dichotomized) predictor. Similar to the ANOVA results, Date did not significantly predict average disgust recall (t = 1.04, p = 0.302). Although such results suggest that COVID-19 may not have had an impact on the results of Experiment 2 over the course of data collection, it is possible that the mere awareness and discussion of the virus resulted in an increase in disgust sensitivity that did not escalate quantitatively until sometime following data collection. Further research on the effects of pandemics and other naturally occurring contamination-based threats on the information processing of disgusting information seems warranted.

The primary limitation of the present investigation is that – contrary to the pilot results reported by Chapman (2018) – the frightening images used were consistently rated as more arousing than the disgusting images used. Such a confound makes it difficult to isolate the effect of interest (i.e. the effect of discrete negative emotions on memory and metamemory). A challenge for researchers in this area will be to develop stimulus sets which eliminate such confounds, and do so consistently across samples.

In light of the present results, a number of future directions for research in this area are suggested. First, given the finding that the LDT may be necessary to observe the disgust advantage in image recall, future research should aim to more formally test the hypothesis that the disgust advantage is stronger under conditions of DA. Such research would allow for a deeper investigation into the mechanisms which mediate the disgust advantage. Additionally, now that it has been determined that the disgust advantage can be replicated using intentional encoding, future research will be better poised to determine whether the disgust advantage extends to metamemorial judgments by adopting methods which resemble those used in Experiment 2.

In conclusion, the present investigation demonstrates that – although the disgust advantage appears to be highly replicable based on the extant literature (Chapman, 2018; Chapman et al., 2013; Charash & McKay, 2002; Croucher et al., 2011; Ferré et al., 2018; Zhang et al., 2018, but see Marchewka et al., 2016) – there are situations in which the effect is unlikely to occur (i.e. under conditions of full attention). In time, other boundary conditions which constrain this effect may appear, shedding further light upon the mechanisms which explain and characterize this effect.

REFERENCES

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, *113*, 518-540. https://psycnet.apa.org/doi/10.1037/0096-3445.113.4.518

Barrett, L. F. (2006). Are emotions natural kinds?. *Perspectives on Psychological Science*, *1*, 28-58. https://doi.org/10.1111%2Fj.1745-6916.2006.00003.x

Begg, I., Vinski, E., Frankovich, L., & Holgate, B. (1991). Generating makes words memorable, but so does effective reading. *Memory & Cognition*, *19*, 487-497. https://doi.org/10.3758/BF03199571

Bennion, K. A., Ford, J. H., Murray, B. D., & Kensinger, E. A. (2013). Oversimplification in the study of emotional memory. *Journal of the International Neuropsychological Society*, *19*, 953-961. https://doi.org/10.1017/S1355617713000945

Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition*, *41*, 897-903. https://doi.org/10.3758/s13421-013- 0307-8

Besken, M., & Mulligan, N. W. (2014). Perceptual fluency, auditory generation, and metamemory: Analyzing the perceptual fluency hypothesis in the auditory modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 429-440. https://doi.org/10.1037/a0034407

Block, R. A. (2009). Intent to remember briefly presented human faces and other pictorial stimuli enhances recognition memory. *Memory & Cognition*, *37*, 667-678. https://doi.org/10.3758/MC.37.5.667

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 379. https://doi.org/10.1037/0278-7393.18.2.379

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin*, *133*, 761-779. https://doi.org/10.1037/0033-2909.133.5.761

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*, 1228-1258. https://doi.org/10.3758/s13415-014-0270-2

Carretié, L., Ruiz-Padial, E., López-Martín, S., & Albert, J. (2011). Decomposing unpleasantness: Differential exogenous attention to disgusting and fearful stimuli. *Biological Psychology*, *86*, 247-253. https://doi.org/10.1016/j.biopsycho.2010.12.005

Chapman, H. A. (2018). Enhanced recall of disgusting relative to frightening photographs is not due to organisation. *Cognition and Emotion*, *32*, 1220-1230. https://doi.org/10.1080/02699931.2017.1394817

Chapman, H. A., Johannes, K., Poppenk, J. L., Moscovitch, M., & Anderson, A. K. (2013). Evidence for the differential salience of disgust and fear in episodic memory. *Journal of Experimental Psychology: General*, *142*, 1100-1112. https://doi.org/10.1037/a0030503

Charash, M., & McKay, D. (2002). Attention bias for disgust. *Journal of Anxiety Disorders, 16*, 529–541. https://doi.org/10.1016/S0887-6185(02)00171-8

Ciesielski, B. G., Armstrong, T., Zald, D. H., & Olatunji, B. O. (2010). Emotion modulation of visual attention: Categorical and temporal characteristics. *PLoS One*, *5*, e13860. https://dx.doi.org/10.1371%2Fjournal.pone.0013860

Cisler, J. M., Olatunji, B. O., Lohr, J. M., & Williams, N. L. (2009). Attentional bias differences between fear and disgust: Implications for the role of disgust in disgust-related anxiety disorders. *Cognition and Emotion*, *23*, 675-687. https://doi.org/10.1080/02699930802051599

Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*, 159-180. https://psycnet.apa.org/doi/10.1037/0096-3445.125.2.159

Croucher, C. J., Calder, A. J., Ramponi, C., Barnard, P. J., & Murphy, F. C. (2011). Disgust enhances the recollection of negative emotional images. *PLoS ONE, 6*, e26571. https://doi.org/10.1371/journal.pone.0026571

D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*, 173-188. https://psycnet.apa.org/doi/10.1037/1528-3542.4.2.173

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*, 169–200. https://doi.org/10.1080/02699939208411068

Ewbank, M. P., Barnard, P. J., Croucher, C. J., Ramponi, C., & Calder, A. J. (2009). The amygdala response to images with impact. *Social Cognitive and Affective Neuroscience*, *4*, 127-133. https://doi.org/10.1093/scan/nsn048

Ferré, P., Haro, J., & Hinojosa, J. A. (2018). Be aware of the rifle but do not forget the stench: Differential effects of fear and disgust on lexical processing and memory. *Cognition and Emotion*, *32*, 796-811. https://doi.org/10.1080/02699931.2017.1356700

Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition and Emotion*, *1*, 115-143. https://doi.org/10.1080/02699938708408043

Girden, E. R. (1992). *ANOVA: Repeated measures*. Newbury Park, CA: Sage.

Hourihan, K. L., & Bursey, E. (2017). A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory*, *25*, 35-43. https://doi.org/10.1080/09658211.2015.1122809

Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory & Cognition*, *45*, 121-136. https://doi.org/10.3758/s13421-016-0646-3

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*, 421-445. https://doi.org/10.1006/jmla.1993.1023

Jay, T., Caldwell-Harris, C., & King, K. (2008). Recalling taboo and nontaboo words. *The American Journal of Psychology*, *121*, 83-103. https://doi.org/10.2307/20445445

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words?. *Memory & Cognition*, *31*, 1169-1180. https://doi.org/10.3758/BF03195800

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, *101*, 3310-3315. https://doi.org/10.1073/pnas.0306408101

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349-370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, *131*, 147-162. https://doi.org/10.1037/0096-3445.131.2.147

Krusemark, E. A., & Li, W. (2011). Do all threats work the same way? Divergent effects of fear and disgust on sensory perception and attention. *Journal of Neuroscience*, *31*, 3429-3434. https://doi.org/10.1523/JNEUROSCI.4394-10.2011

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition and Emotion*, *23*, 833–875. https://doi.org/10.1080/02699930902738863

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, *22*, 530-554. https://doi.org/10.1521/soco.22.5.530.50767

Marchewka, A., Wypych, M., Michałowski, J. M., Sińczuk, M., Wordecha, M., Jednoróg, K., & Nowicka, A. (2016). What is the effect of basic emotions on directed forgetting? Investigating the role of basic emotions in memory. *Frontiers in Human Neuroscience*, *10,* 378. https://doi.org/10.3389/fnhum.2016.00378

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs?. *Memory & Cognition*, *29,* 222-233. https://doi.org/10.3758/BF03194916

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.*, *27*, 1-28. https://doi.org/10.1146/annurev.neuro.27.070203.144157

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*, 159-163. https://doi.org/10.1111/j.1467-8721.2009.01628.x

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*, 174-179. https://doi.org/10.3758/PBR.15.1.174

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*, 61-64.

Murphy, F. C., Hill, E. L., Ramponi, C., Calder, A. J., & Barnard, P. J. (2010). Paying attention to emotional images with impact. *Emotion*, *10*, 605-614. https://psycnet.apa.org/doi/10.1037/a0019681

Nairne, J. S. (1988). The mnemonic value of perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 248-255. https://doi.org/10.1037/0278-7393.14.2.248

Naveh-Benjamin, M., Guez, J., Hara, Y., Brubaker, M. S., & Lowenschuss-Erlich, I. (2014). The effects of divided attention on encoding processes under incidental and intentional learning instructions: Underlying mechanisms?. *The Quarterly Journal of Experimental Psychology*, *67*, 1682-1696. https://doi.org/10.1080/17470218.2013.867517

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109-133. https://doi.org/10.1037/0033-2909.95.1.109

Nomi, J. S., Rhodes, M. G., & Cleary, A. M. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition & Emotion*, *27*, 141-149. https://doi.org/10.1080/02699931.2012.679917

Preuß, D., Schoofs, D., & Wolf, O. T. (2009). Associations between endogenous cortisol levels and emotional memory in young women: Influence of encoding instructions. *Stress*, *12*, 379-387. https://doi.org/10.1080/10253890802524592

Rabinowitz, J. C., Ackerman, B. P., Craik, F. I., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*, *37*, 688-695. https://doi.org/10.1093/geronj/37.6.688

Ramponi, C., Murphy, F. C., Calder, A. J., & Barnard, P. J. (2010). Recognition memory for pictorial material in subclinical depression. *Acta Psychologica*, *135*, 293-301. https://doi.org/10.1016/j.actpsy.2010.07.015

Ritchey, M., Bessette-Symons, B., Hayes, S. M., & Cabeza, R. (2011). Emotion processing in the aging brain is modulated by semantic elaboration. *Neuropsychologia*, *49*, 640-650. https://dx.doi.org/10.1016%2Fj.neuropsychologia.2010.09.009

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*, 615-625. https://doi.org/10.1037/a0013684

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*, 550-554. https://doi.org/10.3758/PBR.16.3.550

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178. https://doi.org/10.1037/h0077714

Schimmack, U., & Derryberry, D. E. (2005). Attentional interference effects of emotional pictures: Threat, negativity, or arousal?. *Emotion*, *5*, 55-66. https://psycnet.apa.org/doi/10.1037/1528-3542.5.1.55

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 294-306. https://doi.org/10.3758/CABN.4.3.294

Shaw, R. J., & Craik, F. I. (1989). Age differences in predictions and performance on a cued recall task. *Psychology and Aging*, *4*, 131-135. https://doi.org/10.1037/0882-7974.4.2.131

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592-604. https://doi.org/10.1037/0278-7393.4.6.592

Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science*, *22*, 430-436. https://doi.org/10.1177/0963721413498893

Talmi, D., Luk, B. T., McGarry, L. M., & Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language*, *56*, 555-574. https://doi.org/10.1016/j.jml.2007.01.002

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, *32*, 742–751. https://doi.org/10.3758/BF03195864

Tauber, S. K., & Dunlosky, J. (2012). Can older adults accurately judge their learning of emotional information?. *Psychology and Aging*, *27*, 924-933. https://doi.org/10.1037/a0028447

Tauber, S. K., Dunlosky, J., Urry, H. L., & Opitz, P. C. (2017). The effects of emotion on younger and older adults' monitoring of learning. *Aging, Neuropsychology, and Cognition*, *24*, 555-574. https://doi.org/10.1080/13825585.2016.1227423

Van Hooff, J. C., Devue, C., Vieweg, P. E., & Theeuwes, J. (2013). Disgust-and not fear-evoking images hold our attention. *Acta Psychologica*, *143*, 1-6. https://doi.org/10.1016/j.actpsy.2013.02.001

Witherby, A. E., & Tauber, S. K. (2018). Monitoring of learning for emotional faces: How do fine-grained categories of emotion influence participants' judgments of learning and beliefs about memory?. *Cognition and Emotion*, *32*, 860-866. https://doi.org/10.1080/02699931.2017.1360252

Zimmer, U., Höfler, M., Koschutnig, K., & Ischebeck, A. (2016). Neuronal interactions in areas of spatial attention reflect avoidance of disgust, but orienting to danger. *NeuroImage*, *134*, 94-104. https://doi.org/10.1016/j.neuroimage.2016.03.050

Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, *62*, 240-253. https://doi.org/10.1016/j.jml.2009.11.004

Zhang, D., Lin, Y., Liu, Y., Luo, Y., & Jiang, D. (2018). Memory encoding, retention and retrieval of disgusting and fearful faces. *Acta Psychologica Sinica*, *51*, 36-47. https://doi.org/10.3724/SP.J.1041.2019.00036

Zhang, D., Liu, Y., Wang, L., Ai, H., & Luo, Y. (2017). Mechanisms for attentional modulation by threatening emotions of fear, anger, and disgust. *Cognitive, Affective, & Behavioral Neuroscience*, *17*, 198-210. https://doi.org/10.3758/s13415-016-0473-9